Surgical Research in Patients: Ideal Time for an IDEAL Checklist

Barkun, Jeffrey S ; Dimick, Justin B ; Clavien, Pierre-Alain

DOI: https://doi.org/10.1097/SLA.0000000000003140

Surgical Research in Patients

Ideal Time for an IDEAL Checklist

Jeffrey S. Barkun, MD, Justin B. Dimick, MD, MPH,† and Pierre-Alain Clavien, MD, PhD‡*

The ability of surgeons to produce and publish high quality clinical research has been challenged on many occasions in past years. Arguably, the most widely remembered can be traced to The Lancet, a journal launched in 1823 by Thomas Wakley:[1] a surgeon intent on improving quality of patient care. In the 1996 editorial entitled ''Surgical research or comic opera: questions, but few answers'',[2] Lancet editor Richard Horton cleverly, albeit bluntly, defied all surgeons to improve the quality of the evidence in our journals. As a rebuke to this challenge, and under the leadership of Jonathan L. Meakins, a series of meetings at Balliol College, Oxford in 2007–2008 brought together surgeons as well as experts in evidence-based medicine, research design, and statistics, the challenge being to reexamine research concepts which had previously been aimed at populational research or the evaluation of medications in the post thalidomide era.[3] The specific goal of the group was to adapt these concepts to the daily, procedure-based surgical reality. The results of these meetings were multifold: the acceptance by methodologists that Surgery is a complex intervention which cannot be evaluated within the existing paradigm of drug treatments and the realization that evaluating surgery involves the integration and documentation of strongly confounding factors such as multiple co-interventions (anaesthesia, physiotherapy, etc…), learning curve effects, the ''iterative process of tinkering'', technical expertise, and challenges of blinding the operator in clinical trials, to name a few.[4]

The IDEAL recommendations and framework were originally proposed in a series of three articles in the Lancet[3–5] in 2009. These describe 5 consecutive stages which correspond to the stages of development and dissemination of a new operative technique. Although they have been described in detail elsewhere,[6] each stage is characterized by posing a specific defining question, as summarized in Table 1 below.

IDEAL proposes stage-specific optimal research designs and outcomes to evaluate the new treatment, ranging from case reports (stage 1) to Randomized Clinical Trials (''RCTs''), and then registries (stage 4). The key evaluation remains a large multi-center, multi surgeon RCT in stage 3, understanding that in some instances, such a RCT may not be feasible or desirable. It should be attempted if at all possible because results are not always predictable, as once again demonstrated by the oncological outcomes of a recent large multi-center RCT comparing minimally invasive to open radical hysterectomy for early-stage cervical cancer, which surprisingly found lower 5-year survival with the minimally invasive approach.[7]

Based on a two-round, on-line delphi process, two broad-based conferences in Oxford in 2016 and New York in 2017 have added to IDEAL, and are reported in this issue of the Annals of Surgery.[8,9] The first addresses a key point to better operationalize IDEAL[10] by adding a PICO system to distinguish which stage most accurately characterizes the new treatment at the time of the evaluation and which endpoints need to be reached prior to moving to the next stage.[8] ''P'' (population) describes how many patients and surgeons (or operators) are currently undergoing/using the new technique, ''I'' (intervention) addresses how mature are the details of the intervention, ''C'' (comparator) speaks to the gold standard which the new technology is trying to replace, and ''O'' (outcomes) are the outcomes, which have already been documented to date and those remaining to document for the current IDEAL stage. In addition, in order to broaden the scope of application of IDEAL, extensions have been proposed: ''Pre-IDEAL'' to address the evaluation of pre-clinical studies prior to ''first in human'' trials; these may include material testing, simulator, cadaver, animal, modelling and cost-effectiveness studies. Other specialized variants include IDEAL-D to evaluate therapeutic devices, or extensions for multiple specific clinical domains such as the R-IDEAL tool from the MRI-Linear Accelerator Consortium[11] and the IDEAL-Physio extension in physiotherapy.[8] In the second report in this issue,[9] ethical considerations are given center stage with a view to

From the *McGill University Health Center, Department of Surgery, Montreal, Canada; †University of Michigan Health Systems, Division of Minimally Invasive Surgery, Ann Arbor, USA; and ‡Department of Surgery and Transplantation, University Hospital Zurich, Zurich, Switzerland.

Disclosure: The authors declare no conflicts of interest.

Reprints: Jeffrey S. Barkun, MD, McGill University Health Center, Department of Surgery, 1001 Décarie boul. Montreal, H4A-3J1, Canada.
E-mail: jeffrey.barkun@muhc.mcgill.ca.

ISSN: 0003-4932/18/26902-0208
DOI: 10.1097/SLA.0000000000003140

TABLE 1. Questions Related to Each Stage of the IDEAL Framework

Stage	Question
Idea Stage 1	What is the new treatment concept and why is it needed?
Development Stage 2a	Has the new intervention reached a state of stability sufficient to allow replication by others?
Exploration Stage 2b	Have the questions that might compromise the chance of conducting a successful Randomized Clinical Trial been addressed?
Assessment Stage 3	How does the new intervention compare with current practice?
Long-term study Stage 4	Are there any long-term or rare adverse effects or changes in indications or delivery quality over time?

minimizing harm to patients, especially during learning curves, all the while respecting patient autonomy and ensuring equitable access to innovations.

Since its introduction a decade ago, there is increasing evidence that the IDEAL approach has been broadly embraced by the surgical community. As of October 2017, a total of 552 papers had cited key IDEAL reports.[8] A paper in the current issue of the Annals of Surgery covers the evolving learning curves in laparoscopic liver surgery according to descriptions of surgeons in stages 2 and 3 of the IDEAL paradigm.[12] Another example is the evaluation of the challenging 2-stage ''ALPPS'' procedure (''associating liver partition and portal vein ligation for staged hepatectomy'') which was described 10 years ago.[13] ALPPS[14] may enable the curative removal of otherwise unresectably large liver tumors, but has been found to be associated with high morbidity. This innovative procedure has also been consistently reported according to the IDEAL stages.[15]

There is increasing acceptance of IDEAL among health care decider communities internationally: the Medical Device Epidemiology Network Initiative (MDEpiNet) partnership with the US Food and Drug Administration (FDA);[16] EXCITE: an international consortium offering device manufacturers a comprehensive assessment service for innovative new products; and the European Network for Health Technology Assessment (EUnetHTA National Institute of Healthcare Research) in the UK which recommends IDEAL study designs in commissioning briefs.[8]

Given the seemingly wide-ranging consensus to accept this framework for the evaluation of new surgical technology, what would be the next steps to continue to meet Richard Horton's challenge head on? The moment certainly appears to be ripe for Surgery. Seen in a broader perspective, the blooming of IDEAL embodies a significant underlying paradigm shift: until the 70's-80's, formal surgical research training and tools consisted essentially of assays and wet lab exposure with the quality of clinical work being assessed primarily through institutional mortality/morbidity rounds,[17] in the tradition of E. Codman,[18] as well as by the publication of case reports and retrospective case series. Other than a few notable cases, surgery initially seemed to be left out of the evidence-based revolution,[19] as shown by the penetration of RCT's in the surgical literature estimated to have been consistently less than 10% in 1995 and 2006.[20,21] Flash forward to 2018 when a random sampling of surgical literature[22] suggests this figure may have tripled or even quadrupled. What's more, even if the quality of published surgical research may not be convincingly on the upswing,[23,24] the potential and the tools to advance it are solidly in place. Unlike their predecessors, mainstream clinical researchers in surgery are currently able to call upon a multitude of validated or consensus-driven tools to address most aspects of performing, measuring, and reporting clinical surgical research. These tools are complementary to the framework proposed by IDEAL and represent the next generation of ''assays'' in modern surgical clinical research. They include generic outcome measures such as the Clavien-Dindo[25,26] classification, or comprehensive complication index (CCI)[27] for morbidity/mortality, and procedure-specific validated measures such as the classification for delayed gastric emptying after pylorus-preserving pancreaticoduodenectomy by the international Study Group of Pancreatic Surgery.[28] The COMET initiative is committed to the development and application of agreed standardized sets of outcomes, known as ''core outcome sets''.[29] We surgeons also have our own surgery-specific, technically-intensive RCT designs such as ''expertise-based trials'' where patients are randomized to a single surgeon performing a single procedure rather than multiple surgeons performing more than one procedure which they may not fully master or believe in.[30] We even have a grid to evaluate if a published RCT truly fits our practice (''Applicability Scoring of Surgical trials'').[31]

The EQUATOR Network promotes transparent and accurate reporting and a wider use of reporting guidelines in biomedical science.[32] Its sponsors keep track of a plethora of reporting standards many of which are fine-tuned for surgical practice such as the CONSORT statement for non-pharmacological randomized studies),[33] which addresses the documentation of variables known to be important in surgery rather than drug trials, especially in stage 3. Examples are *''details of the intervention and comparator as they were implemented'', and ''a description of care providers; case volume, qualification, expertise and centers volume in each group. . .''*. The (TIDieR) checklist and guide (''template for intervention description and replication''[34] is committed to the better reporting of interventions by describing these in sufficient detail to allow their replication, essential in IDEAL stages 2 and 2a. The SCARE statement and checklist applies to surgical case reports such as are needed in stages 1 and 2 of IDEAL.[35] EQUATOR also tracks proposed reporting guidelines for newer methodologies which are increasingly popular in surgery such as the proposed Reporting and Guidelines in Propensity Score Analysis used in Cancer Surgical Studies,[36] or CONSORT and STROBE statement extensions in simulation research.[37]

With all these surgery-specific pieces in place, what can the Annals of Surgery do to help its readership bring our literature to the next level?

A strategy would be for journal editors to make the IDEAL framework a mandatory part of the submission process. Authors submitting publications that evaluate new surgical procedures and technology could use a series of checklists to identify the IDEAL stage and then note the presence or absence of important design features in the conduct and reporting of their study. There is a precedent for this type of structured checklist assessment as part of the submission process. The CONSORT guidelines for reporting

randomized clinical trials have been widely used.[38,39] The checklist is part of the instructions for authors and completing the checklist is mandatory at many journals.

Peer reviewers have access to this checklist to compare the design features of the trial with recommended best practices. Although it is difficult to measure the impact of this checklist in improving the quality of clinical trials, there is compelling evidence that it improves the quality of reporting key elements of design.[40]

To implement this strategy for the IDEAL framework, specific checklists of best practices in conduct and reporting would still need to be developed for each IDEAL stage. This editorial should thus be seen as a not-so-cleverly disguised challenge to develop such a set of IDEAL checklists in order to permit us to definitively write the last act of our surgical comic opera.

REFERENCES

1. Thomas Wakley. Available at: https://en.wikipedia.org/wiki/Thomas_Wakley. Last accessed 11/13/2018.
2. Horton R. Surgical research or comic opera: questions, but few answers. *Lancet*. 1996;347:984–985.
3. Barkun JS, Aronson JK, Feldman LS, et al. Evaluation and stages of surgical innovations. Balliol Collaboration. *Lancet*. 2009;374:1089–1096.
4. McCulloch P, Altman DG, Campbell WB, et al. No surgical innovation without evaluation: the IDEAL recommendations. Balliol Collaboration. *Lancet*. 2009;374:1105–1112.
5. Ergina PL, Cook JA, Blazeby JM, Boutron I, et al. Challenges in evaluating surgical innovation. Balliol Collaboration. *Lancet*. 2009;374:1097–1104.
6. McCulloch P, Cook JA, Altman DG, et al. IDEAL framework for surgical innovation 1: the idea and development stages. *BMJ*. 2013;346:f3012.
7. Ramirez PT, Frumovitz M, Pareja R, et al. Minimally Invasive versus Abdominal Radical Hysterectomy for Cervical Cancer. *N Engl J Med*. 2018 [Epub ahead of print].
8. Hirst A, Philippou Y, Blazeby J, et al. No surgical innovation without evaluation: Evolution and further development of the IDEAL Framework and Recommendations. *Ann Surg*. 2019;269:211–220.
9. Rogers WA, Hutchison K, McNair A. Ethical issues across the IDEAL stages of surgical innovation. *Ann Surg*. 2019;269:229–233.
10. Pennell CP, Hirst AD, Campbell WB, et al. Practical guide to the Idea, Development and Exploration stages of the IDEAL Framework and Recommendations. *British J Surgery*. 2016;103:607–615.
11. Verkooijen HM, Kerkmeijer LG, Fuller CD, et al. R-IDEAL: A Framework for Systematic Clinical Evaluation of Technical Innovations in Radiation Oncology. *Front Oncol*. 2017;7:59.
12. Halls CM, Alseidi A, Berardi G, et al. A comparison of the learning curves of laparoscopic liver surgeons in differing stages of the IDEAL paradigm of surgical innovation: standing on the shoulders of pioneers. *Ann Surg*. 2019;269:221–228.
13. Schnitzbauer AA, Lang SA, Goessmann H, et al. Right portal vein ligation combined with in situ splitting induces rapid left lateral liver lobe hypertrophy enabling 2-staged extended right hepatic resection in small-for-size settings. *Ann Surg*. 2012;255:405–414.
14. de Santibañes E, Clavien PA. Playing Play-Doh to prevent postoperative liver failure: the "ALPPS" approach. *Ann Surg*. 2012;255:415–417.
15. Clavien PA, Lillemoe KD. Associating Liver Partition and Portal Vein Ligation for Staged Hepatectomy. *Ann Surg*. 2016;263:835–836.
16. The Medical Device Epidemiology Network Initiative (MDEpiNet) partnership with the US Food and Drug Administration (FDA). Available at: http://www.ideal-collaboration.net/2014/03/joint-ideal-cornell-and-fda-meeting-to-discuss-device-regulation-in-new-york-3rd-4th-april-2014.
17. Feldman L, Barkun J, Barkun A, et al. Measuring postoperative complications in general surgery patients using an outcomes-based strategy: comparison with complications presented at morbidity and mortality rounds. *Surgery*. 1997; 122:711–719.
18. Ernest Amory Codman. Available at: https://en.wikipedia.org/wiki/Ernest_Amory_Codman. Last accessed 13/11/2018.
19. Black N. Evidence-based surgery: a passing fad? *World J Surgery*. 1999; 23:789–793.
20. Solomon MJ, McLeod RS. Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery*. 1995;118:459–467.
21. Chang DC, Matsen SL, Simpkins CE. Why should surgeons care about clinical research methodology? *J Am Coll Surg*. 2006;203:827–830.
22. McCulloch P, Feinberg J, Philippou Y, et al. Progress in clinical research in surgery and IDEAL. *Lancet*. 2018;392:88–94.
23. Hannink G, Gooszen HG, Rovers MM. Comparison of Registered and Published Primary Outcomes in Randomized Clinical Trials of Surgical Interventions. *Ann Surg*. 2013;257:818–823.
24. Agha R, Cooper D, Muir G. The reporting quality of randomised controlled trials in surgery: A systematic review. *International J Surgery*. 2007;5:413–422.
25. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240:205–213.
26. Clavien PA, Barkun JS, de Oliveira. et al. The Clavien-Dindo classification of surgical complications: five-year experience. *Ann Surg*. 2009;250:187–196.
27. Slankamenac K, Graf R, Barkun JS, et al. The Comprehensive Complication Index (CCI) A Novel Continuous Scale to Measure Surgical Morbidity. *Ann Surg*. 2013;258:1–7.
28. Giuseppe Malleo, Stefano Crippa, Giovanni Butturini. et al. Delayed gastric emptying after pylorus-preserving pancreaticoduodenectomy: validation of International Study Group of Pancreatic Surgery classification and analysis of risk factors. *HPB*. 2010;12:610–618.
29. COMET intiative. Available at: http://www.comet-initiative.org/ last accessed on November 12th, 2018.
30. Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ*. 2005;330:88.
31. Tourabaly I, Boutron I, Nizard R, et al. ASSIST Applicability Scoring of Surgical trials. An Investigator-reported aSsessment Tool. *PLoS ONE*. 2012;7:e42258.
32. EQUATOR Network. Available at: http://www.equator-network.org.
33. Boutron I, Altman DG, Moher D, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. CONSORT NPT Group. *Ann Intern Med*. 2017;167:40–47.
34. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687.
35. Agha RA, Fowler AJ, Saeta A, et al. The SCARE Statement: Consensus-based surgical case report guidelines. SCARE Group. *Int J Surg*. 2016;34:180–186.
36. Yao XI, Wang X, Speicher PJ, et al. Reporting and Guidelines in Propensity Score Analysis: A Systematic Review of Cancer and Cancer Surgical Studies. *J Natl Cancer Inst*. 2017;109.
37. Cheng A, Kessler D, Mackinnon R, et al. International Network for Simulation-based Pediatric Innovation, Research, and Education (INSPIRE) Reporting Guidelines Investigators. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simul Healthc*. 2016;11:238–248.
38. Schulz KF, Altman DG, Moher D, et al., for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Lancet*. 2010;375:1136.
39. Clavien PA, Lillemoe KD. A new policy to implement CONSORT guidelines for surgical randomized controlled trials. *Ann Surg*. 2014;260:947–948.
40. Moher D, Jones A, Lepage L, et al., CONSORT Group. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and after evaluation. *JAMA*. 2001;285:1992–1995.